above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198 (1998), and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (e.g., vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch *et al.*, *Proc. Natl. Acad. Sci. USA* 86:317-321 (1989); Flexner *et al.*, *Ann. N.Y. Acad. Sci.* 569:86-103 (1989); Flexner *et al.*, *Vaccine* 8:17-21 (1990); U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627 (1988); Rosenfeld *et al.*, *Science* 252:431-434 (1991); Kolls *et al.*, *Proc. Natl. Acad. Sci. USA* 91:215-219 (1994); Kass-Eisler *et al.*, *Proc. Natl. Acad. Sci. USA* 90:11498-11502 (1993); Guzman *et al.*, *Circulation* 88:2838-2848 (1993); and Guzman *et al.*, *Cir. Res.* 73:1202-1207 (1993). Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer *et al.*, *Science* 259:1745-1749 (1993) and reviewed by Cohen, *Science* 259:1691-1692 (1993). The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells. It will be apparent that a vaccine may comprise both a polynucleotide and a polypeptide component. Such vaccines may provide for an enhanced immune response.

It will be apparent that a vaccine may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (e.g., salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (e.g., sodium, potassium, lithium, ammonium, calcium and magnesium salts).

While any suitable carrier known to those of ordinary skill in the art may be employed in the vaccine compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be

formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (e.g., polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. One may also employ a carrier comprising the particulate-protein complexes described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (e.g., neutral buffered saline or phosphate buffered saline), carbohydrates (e.g., glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (e.g., aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of immunostimulants may be employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bortadella pertussis* or *Mycobacterium* species or *Mycobacterium* derived proteins. For example, delipidated, deglycolipidated *M. vaccae* ("pVac") can be used. In another embodiment, BCG is used as an adjuvant. In addition, the vaccine can be administered to a subject previously exposed to BCG. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 and derivatives thereof (SmithKline Beecham, Philadelphia, PA); CWS, TDM, Leif, aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars;

cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN-γ, TNFα, IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann & Coffman, *Ann. Rev. Immunol.* 7:145-173 (1989).

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato *et al.*, *Science* 273:352 (1996). Another preferred adjuvant comprises a saponin, such as Quil A, or derivatives thereof, including QS21 and QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; or *Gypsophila* or *Chenopodium quinoa* saponins . Other preferred formulations include more than one saponin in the adjuvant combinations of the present invention, for example combinations of at least two of the following group comprising QS21, QS7, Quil A, β-escin, or digitonin.

Alternatively the saponin formulations may be combined with vaccine vehicles composed of chitosan or other polycationic polymers, polylactide and polylactide-co-glycolide particles, poly-N-acetyl glucosamine-based polymer matrix,

particles composed of polysaccharides or chemically modified polysaccharides, liposomes and lipid-based particles, particles composed of glycerol monoesters, etc. The saponins may also be formulated in the presence of cholesterol to form particulate structures such as liposomes or ISCOMs. Furthermore, the saponins may be formulated together with a polyoxyethylene ether or ester, in either a non-particulate solution or suspension, or in a particulate structure such as a paucilamelar liposome or ISCOM. The saponins may also be formulated with excipients such as Carbopol® to increase viscosity, or may be formulated in a dry powder form with a powder excipient such as lactose.

In one preferred embodiment, the adjuvant system includes the combination of a monophosphoryl lipid A and a saponin derivative, such as the combination of QS21 and 3D-MPL® adjuvant, as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. Another particularly preferred adjuvant formulation employing QS21, 3D-MPL® adjuvant and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Another enhanced adjuvant system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG and QS21 as disclosed in WO 00/09159. Preferably the formulation additionally comprises an oil in water emulsion and tocopherol.

Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2, AS2', AS2,'' SBAS-4, or SBAS6, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties, and polyoxyethylene ether adjuvants such as those described in WO 99/52549A1.

Other preferred adjuvants include adjuvant molecules of the general formula (I): $HO(CH_2CH_2O)_n$-A-R,

wherein, n is 1-50, A is a bond or –C(O)-, R is C_{1-50} alkyl or Phenyl C_{1-50} alkyl.

One embodiment of the present invention consists of a vaccine formulation comprising a polyoxyethylene ether of general formula (I), wherein *n* is between 1 and 50, preferably 4-24, most preferably 9; the *R* component is C_{1-50}, preferably C_4-C_{20} alkyl

and most preferably C_{12} alkyl, and *A* is a bond. The concentration of the polyoxyethylene ethers should be in the range 0.1-20%, preferably from 0.1-10%, and most preferably in the range 0.1-1%. Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether, polyoxyethylene-9-steoryl ether, polyoxyethylene-8-steoryl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether. Polyoxyethylene ethers such as polyoxyethylene lauryl ether are described in the Merck index (12th edition: entry 7717). These adjuvant molecules are described in WO 99/52549.

The polyoxyethylene ether according to the general formula (I) above may, if desired, be combined with another adjuvant. For example, a preferred adjuvant combination is preferably with CpG as described in the pending UK patent application GB 9820956.2.

Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (i.e., a formulation such as a capsule, sponge or gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (*see, e.g.*, Coombes *et al.*, *Vaccine* 14:1429-1438 (1996)) and administered by, for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (e.g., a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see, e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (i.e., matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau & Steinman, *Nature* 392:245-251 (1998)) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman & Levy, *Ann. Rev. Med.* 50:507-529 (1999)). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel *et al.*, *Nature Med.* 4:594-600 (1998)).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNFα to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNFα, CD40 ligand, LPS, flt3 ligand and/or

other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (e.g., CD54 and CD11) and costimulatory molecules (e.g., CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a protein (or portion or other variant thereof) such that the polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi *et al.*, *Immunology and Cell Biology* 75:456-460 (1997). Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (e.g., vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (e.g., a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

Vaccines and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or

aqueous vehicles. Alternatively, a vaccine or pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a protein of the invention.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

EXAMPLES

The following examples are provided by way of illustration only and not by way of limitation. Those of skill in the art will readily recognize a variety of noncritical parameters that could be changed or modified to yield essentially similar results.

Example 1: Guinea pig vaccination with MTB72F fusion protein and compositions with individual antigens

Guinea pigs were immunized with adjuvant alone (SBAS1, SBAS2, or ASAS7 plus Al(OH)3), MTB72F fusion protein in adjuvant, or TbH9 plus Ra35 antigen composition.

Methods:

Groups:
1) SBAS1
2) SBAS2
3) SBAS7 + Al(OH)3
4) TbH9+Ra35 + SBAS1
5) TbH9 + Ra35 + SBAS2
6) TbH9 + Ra35 + SBAS7(Al(OH)3)
7) MTB72F in SBAS1
8) MTB72F in SBAS2
9) MTB72F in SBAS7+Al(OH)3
10) PBS
11) BCG

Dosage: 4 μg each of TbH9 and Ra35
8 μg MTB72F

Protocol: 1st immunization, 2nd immunization approximately 3 weeks later, 3rd immunization approximately two and a half weeks later.

Pre-challenge: DTH (delayed type hypersensitivity, used to determine antigenicity; 10 μg antigen)

Challenge: Aerosol with ~30 cfu Erdman strain

Post challenge monitoring: Weight loss
Death (~6 months post challenge)

Results:

1. DTH

Positive reaction to the immunizing antigens. Reactions to individual antigens or the fusion protein were comparable. Skin test reactivity to PPD was only seen with the BCG immunized groups

2. Protection: Guinea pigs vaccinated with MTB72F fusion protein afforded protection compared to those immunized with a mixture of antigens (*see* Figure 1).

Example 2: Mouse vaccination with MTB72F fusion protein and compositions with individual antigens

As described above, mice were immunized with adjuvant alone (SBAS2, SBAS2', SBAS2'', or SBAS6), MTB72F fusion protein in adjuvant, MTB72F DNA, MTB59F fusion protein in adjuvant, or TbH9, Ra35 and Ra12 antigen composition.

Methods:

Groups:
1) MTB72F+ SBAS2
2) MTB72F + SBAS2'
3) MTB72F + SBAS2''
4) MTB72F + SBAS6
5) Ra12+ TbH9 + Ra35 in SBAS2
6) MTB59F in SBAS2
7) SBAS2
8) MTB72F + delipidated, deglycolipidated *M. vaccae*
9) MTB72F DNA
10) MTB72F +IFA
11) MTB72F + BCG

12) delipidated, deglycolipidated *M. vaccae*
13) BCG
14) Saline
15) MTB72F +SBAS2 (in house formulation)

8 animals per group

Immunization schedule: First immunization, second immunization approximately 3 weeks later; third immunization approximately three weeks later.

Aerosol challenge approximately three months after first does

Spleen or lung cells were isolated and cultured; count CFU of cultures approximately three weeks after plating.

Dose: 8 μg MTB72F, 6.56 μg MTB59F, or 1.52, 4.3, and 2.24 μg, respectively, of Ra12, TbH9, and Ra35, mixed.

Results:

Of the AS adjuvants, AS2" + MTB72F gave the best protection in both the spleen and lung in this set of experiments (*see* Figures 2A and 2B). MTB72F gave ~1 log better protection than MTB59F in both spleen and lung in this set of experiments, indicating that Ra12 provides additional benefit. Mixture of 12/H9/35 + AS2 gave a better protection than MTB72F in this experiment. MTB72F DNA gave the best protection in this experiment, particularly in the spleen (>2 log). The protection was comparable in the lung to that seen with MTB72F protein + AS2", in this experiment.

Example 3: Guinea pig vaccination with MTB72F fusion protein and compositions with individual antigens

As described above, guinea pigs were immunized with adjuvant alone (SBAS2, SBAS2', SBAS2'', or SBAS6), MTB72F fusion protein in adjuvant, MTB72F DNA, MTB59F fusion protein in adjuvant, or TbH9, Ra35 and Ra12 antigen composition.

Methods:

Groups: 1) MTB72F + SBAS2
2) MTB72F + SBAS2'

3) MTB72F + SBAS2''
4) MTB72F + SBAS6
5) Ra12+ TbH9 + Ra35 in SBAS2
6) MTB59F in SBAS2
7) SBAS2
8) MTB72F + pvac
9) MTB72F DNA
10) MTB72F +IFA
11) MTB72F + BCG
12) BCG
13) Saline
14) delipidated, deglycolipidated *M. vaccae*

Antigens:

Antigens were formulated on a molar equivalent

5 animals per group

Injection volume per dose is 250µl (IM) containing

MTB72F	20 µg
Ra12, TbH9, Ra35	3.8, 10.8, and 5.6 µg
MTB59F	16.4 µg

Schedule:

1st immunization, 2nd immunization approximately three weeks later, 3rd immunization approximately three weeks later.

Challenge: ~ one and one half months after first immunization.

Results:

~38 Wks post challenge

Groups	Alive	State
G1. MTB72F + AS2	1/5	[losing weight]

G2. MTB72F + AS2'	2/5	[not gaining weight]
G3. MTB72F + AS2"	3/5	[looking okay, but no weight gain]
G4. MTB72F + AS6	2/5	[both those gaining weight]
G5. MTBRa12+H9+Ra35 +AS2	4/5	[one maybe a bit peaked, but two gaining]
G6. MTB59F + AS2	2/5	[both losing a little]
G7. AS2	2/5	[both losing]
G8. MTB72F + pVac	1/5	[not looking too good]
G9. MTB72F DNA	3/5	[all holding steady]
G10. MTB72F + IFA	2/5	[doing okay]
G11. MTB72F + BCG	5/5	[eating very well]
G12 BCG	4/5	[doing fine]
G13 Saline	all dead	
G14 pVac	2/5	[not gaining weight]

By 50 weeks post challenge, while 80% (4/5) of the guinea pigs immunized with BCG + Mtb72F were still alive, only 20% (1/5) of those immunized with BCG alone were alive. At 85 weeks, 4/5 of the guinea pigs immunized with BCG + Mtb72F were still alive and healthy (*see* Figure 7).

Example 4: Long term protection

As described above, guinea pigs were immunized with adjuvant alone (AS2 or AS2''), MTB72F fusion protein in adjuvant, TbH9, Ra35 and Ra12 antigen composition, or a variety of individual antigens in adjuvant.

Methods

GROUPS	ANTIGEN DOSE
1. AS2" + MTB39 (TbH9)	20ug/250ul (IM)
2. AS2" + MTB8.4 (DPV)	20ug
3. AS2" + MTB9.9 (MTI)	20ug
4. AS2" + MTB41 (MTCC#2)	20ug
5. AS2" + MTB40 (HTCC#1)	20ug
6. AS2" + MTB9.8 (MSL)	20ug
7. AS2" + MTB72F	20ug

8. AS2" + Ra12+TbH9 + Ra35 (molar equivalent) 3.8 μg +10.8 μg +5.6 μg
9. AS2" + MTB71F + MTB72F+HTCC#1 20 μg +20 μg +10 μg
10. AS2" + Ra12 20 μg
11. BCG
12. AS2"
13. AS2 + MTB72F
14. AS2+ Rs12+TbH9+Ra35
15. AS2

Example 5: Monkey vaccination with MTB72F fusion protein and compositions with individual antigens

As described above, monkeys were immunized with MTB72F fusion protein in SBAS2 adjuvant, or MTB8.4 antigen composition in adjuvant, or a mixture of MTB72F and MTB8.4.

Methods:

Groups

1. Saline
2. BCG
3. MTB8.4/AS2
4. MTB72F/AS2
5. MTB72F/AS2 (one arm) + MTB8.4/AS2 (other arm)

40 μg each antigen

Results:

At 8 weeks post challenge, monkeys immunized with BCG are showing signs of infection

Current data for 16 weeks post challenge reveals the following trend:

Groups immunized with MTB72F (4 and 5) are holding on their weights and have low ESR values compared to group 3 (MTB8.4 immunization) (Tables 1 and 2).

Table 1

Prophylactic Vaccine Study in Cynomolgus Monkeys with MTB8.4 and MTB72F formulated in AS2 20 Weeks Post Challenge

Groups	ID	Net weight Change (kg)	Chest X-ray (onset)	Status
	1398K	-24%	Pn, bil, prog (wk 8)	Alive
AS2	4437B	-33%	Pn, bil, prog (wk4)	Dead
	2959G	-8.30%	Pn, bil, prog (wk4)	Alive
	605AE	-14.00%	Pn, rt, stable (wk 8)	Alive
	3436A	-15.00%	Neg	Alive
BCG	3642G	Plus 4.5%	Pn, rt, prog (wk 8)	Alive
	1190H	0%	Neg	Alive
	1051I	-30%	Pn, rt, prog (wk 8)	Dead
	3665C	-25%	Pn, rt, prog (wk8)	Dead
MTB8.4	2200F	-18.00%	Pn, rt, stable (wk8)	Alive
	1654J	-33.00%	Pn, bil, prog (wk4)	Dead
	4141C	-33%	Pn, bil, prog (wk4)	Dead
	3061C*	Died after IT challenge		
MTB72F	1228G	Plus 3.6%	Bron, bil, stable for 3 mo (wk8)	Alive
	3462E	-2.20%	Neg	Alive
	4254C	Plus 1.21	Pn, rt, stable for 3 mo (wk4)	Alive
	4496A	Plus 7%	Pn, rt, stable for 1 mo (wk 8)	Alive
MTB8.4	4422C	-39.00%	Pn, bil, prog (wk 4)	Dead
MTB72F	4416A	Plus 11%	Pn, rt, stable for 2 mo (wk 12)	Alive
	2734E	Plus 12.5%	Susp infil rt, stable for 3 mo (wk 8)	Alive

Table 2

Prophylactic Vaccine Study in Cynomolgus Monkeys with MTB8.4 and MTB72F formulated in AS2

		Wks Post Challenge ESR				
Groups	ID	4	8	12	16	16 wks Chest X-r
	1398K	3	3	10	19	Pn, bil, progrsv
AS2	4437B	10	20	3		Died
	2959G	6	3	3	0	Pn, rt, progrsv
	605AE	1	4	7	3	Pn, rt, stable
	3436A	0	8	7	15	Neg
BCG	3642G	0	0	0	0	Pn, rt, progrsv
	1190H	1	0	2	0	Neg
	1051I	0	8	22	7	Pn, bil, w/furt pro
						Died
	3665C	12	30	19		Died
MTB8.4	2200F	1	7	2	0	Pn, rt, progrsv
	1654J	20	8	21	7	Pn,bil,w/fur progr
	4141C	13	8	2	15	Pn,bil,w/fur progr
	3061C*	Died after IT challenge				
MTB72F	1228G	0	1	20	0	Now stable
	3462E	0	0	0	0	Neg
	4254C	13	0	0	0	Pn, now stable
	4496A	5	1	0	5	Pn, rt, w/furt prog
MTB8.4/	4422C	10	3	0		Died
MTB72F	4416A	6	0	1	0	Pn, now stable
	2734E	0	0	0	0	Susp infil, now st

Example 6: BCG priming experiment in monkeys

5 animals per group with four groups immunized with BCG and then rested, then immunized as described above and challenged. The following protocol will be used:

Groups	# animals	Immunizing Antigen	Antigen Dose
1. Nothing	5	AS2	
2. BCG	5	AS2	
3. BCG	5	MTB72F	40ug
4. BCG	4	Ra12+TbH9+Ra35	Molar equiv of antigens in MTB72F dose
5. BCG	4	MTB72F + MTB71F + MTB40	40ug MTB72F 40ug MTB72F 20ug MTB40

All antigens in formulated in AS2

Groups 4 and 5 have four animals each. Two of the BCG immunized monkeys died

Groups	# animals	Immunizing Antigen	Antigens for T cell proliferation and cytokine production assays
1. Nothing	5	AS2	PHA, PPD, MTB72F, MTB71F, HTCC#1, DPV, MTCC#2, Ra12, TbH9, Ra35, MSL, MTI
2. BCG	5	AS2	PHA, PPD, MTB72F, MTB71F, HTCC#1, DPV, MTCC#2, Ra12, TbH9, Ra35, MSL, MTI
3. BCG	5	MTB72F	PHA, PPD, MTB72F, Ra12, TbH9, Ra35
4. BCG	4	Ra12+TbH9+Ra35	PHA, PPD, MTB72F, Ra12, TbH9, Ra35
5. BCG	4	MTB72F + MTB71F + MTB40	PHA, PPD, MTB72F, MTB71F, HTCC#1, DPV, MTCC-2, Ra12, TbH9, Ra35, MSL, MTI

Example 7: Construction of Ra35MutSA and MTB72FMutSA

Expression of Mtb72f typically results in some breakdown products. In addition, the expression of the full-length sequences of the mature or full length form of Ra35 (Mtb32A) in *E. coli* has been difficult. The expressed product was only visible after immunoblotting with a polyclonal rabbit anti-Ra35 Ab indicative of low levels of protein expression. Even then, multiple specific species (bands) were detected indicative of auto-catalytic breakdown (degradation) of the recombinant antigen. This was presumed to be due to the expression of Ra35FL in *E. coli* as a biologically active form.

It has been previously shown that it was possible to express Ra35FL as two overlapping halves comprising the N-terminal (Ra35N-term, called Ra35) and C-term halves (Ra35C-term called Ra12). To enhance and stabilize the expression of the whole Ra35 molecule, a single point mutation was introduced at one of the residues

within the active-site triad (substitution of Ser to Ala; *see* Figures 6). This mutagenized form of Mtb32A can now be easily expressed at high levels in a stable form. In addition, to stabilize expression of Mtb72F, a single nucleotide substitution (T to G, resulting in a Ser to Ala change at position 710 of the fusion polypeptide) was incorporated in the sequence of Mtb72F at nucleotide position 2128 (*see* Figure 5).

This stabilization is also readily accomplished by mutagenizing any one, any two, or all three of the three residues comprising the active site triad in Ra35FL, Ra35, or Mtb72F or other fusion proteins comprising Ra35 (His, Asp, or Ser). Mutagenesis can be performed using any technique known to one of skill in the art.

Example 8: Immunization of mice withf Ra35FLMutSA-TbH9 and MTB72FMutSA

Eight mice per group were immunized with the compositions listed below, which include the adjuvant AS2A. The mice were then challenged with *Mycobacterium tuberculosis*, and survival of the mice was measured.

Group	Concentration of protein or DNA
1. Mtb72f protein	1.5 mg/ml
2. Mtb72f DNA	1.2 mg/ml
3. Mtb72f-85b protein	0.6 mg/ml
4. Mtb72f-85b DNA	1.1 mg/ml
5. Mtb72f-MTI protein	1.3 mg/ml
6. Mtb72f-MTI DNA	1.1 mg/ml
7. Mtb72f MutSA protein	1.7 mg/ml
8. MTB3AMutSA-TbH9 protein	2.4 mg/ml
9. BCG	
10. AS2	
11. vector alone	1.5 mg/ml
12. saline	

WHAT IS CLAIMED IS

1. A composition comprising a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex, and a MTB32A antigen (SEQ ID NO:2 or 4) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex.

2. The composition of claim 1, comprising a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex, and a polypeptide comprising at least 195 amino acids from the N-terminus of a MTB32A antigen (SEQ ID NO:2 or 4) from a *Mycobacterium* species of the tuberculosis complex.

3. The composition of claim 2, further comprising a polypeptide comprising at least about 132 amino acids from the C-terminus of MTB32A antigen (SEQ ID NO:2 or 4) from a *Mycobacterium* species of the tuberculosis complex.

4. The composition of claims 1, 2, or 3, wherein the antigens are covalently linked, thereby forming a fusion polypeptide.

5. The composition of claim 4, wherein the fusion polypeptide has the amino acid sequence of MTB59F (SEQ ID NO:20).

6. The composition of claim 4, wherein the fusion polypeptide has the amino acid sequence of MTB72F (SEQ ID NO:16).

7. The composition of claim 4, wherein the fusion polypeptide has the amino acid sequence of MTB72FMutSA (SEQ ID NO:18).

8. The composition of claim 6 or 7, further comprising BCG.

9. The composition of claim 6 or 7, further comprising at least one additional antigen from a *Mycobacterium* species of the tuberculosis complex, wherein the antigen is selected from the group consisting of MTB8.4 antigen (SEQ ID NO:22), MTB9.8 antigen (SEQ ID NO:24), MTB9.9 antigen (SEQ ID NO:27), MTB40 antigen (SEQ ID NO:29), MTB41 antigen (SEQ ID NO:31), 38-1 (SEQ ID NO:35), TbRa3 (SEQ ID NO:37), 38 kD (SEQ ID NO:39), DPEP (SEQ ID NO:41), TbH4 (SEQ ID NO:43),

DPPD(SEQ ID NO:45), MTB82, Erd14, ESAT-6 antigen (SEQ ID NO:33), MTB85 complex antigen, or α-crystalline antigen, or an immunogenic fragment thereof.

10. The composition of claim 6 or 7, further comprising an adjuvant.

11. The composition of claim 4, wherein the antigens are covalently linked via a chemical linker.

12. The composition of claim 11, wherein the chemical linker is an amino acid linker.

13. The composition of claim 1, further comprising at least one additional antigen from a *Mycobacterium* species of the tuberculosis complex, wherein the antigen is selected from the group consisting ofMTB8.4 antigen (SEQ ID NO:22), MTB9.8 antigen (SEQ ID NO:24), MTB9.9 antigen (SEQ ID NO:27), MTB40 antigen (SEQ ID NO:29), MTB41 antigen (SEQ ID NO:31), 38-1 (SEQ ID NO:35), TbRa3 (SEQ ID NO:37), 38 kD (SEQ ID NO:39), DPEP (SEQ ID NO:41), TbH4 (SEQ ID NO:43), DPPD(SEQ ID NO:45), MTB82, Erd14, ESAT-6 antigen (SEQ ID NO:33), MTB85 complex antigen, or α-crystalline antigen, or an immunogenic fragment thereof.

14. The composition of claim 1, further comprising an adjuvant.

15. The composition of claim 14, wherein the adjuvant comprises QS21 and MPL.

16. The composition of claim 14, wherein the adjuvant is selected from the group consisting of AS2, ENHANZYN, MPL, 3D-MPL, IFA, QS21, CWS, TDM, AGP, CPG, Leif, saponin, and saponin mimetics.

17. The composition of claim 1, further comprising BCG or pVac.

18. The composition of claim 1, further comprising an NS1 antigen or an immunogenic fragment thereof.

19. The composition of claim 1, wherein the *Mycobacterium* species is *Mycobacterium tuberculosis*.

20. An expression cassette comprising a nucleic acid encoding a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex, and a nucleic acid encoding a MTB32A antigen (SEQ ID NO:2 or 4) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex.

21. The expression cassette of claim 20, comprising a nucleic acid encoding a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex, and a nucleic acid encoding a polypeptide comprising at least 195 amino acids from the N-terminus of a MTB32A antigen (SEQ ID NO: 2 or 4) from a *Mycobacterium* species of the tuberculosis complex.

22. The expression cassette of claim 21, further comprising a nucleic acid encoding a polypeptide comprising at least 132 amino acids of the C-terminus of a MTB32A antigen (SEQ ID NO:2 or 4) from a *Mycobacterium* species of the tuberculosis complex.

23. The expression cassette of claim 20, wherein the nucleic acid encodes a fusion polypeptide comprising a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof and a nucleic acid encoding a MTB32A antigen (SEQ ID NO:2 or 4) or an immunogenic fragment thereof.

24. The expression cassette of claim 23, wherein the nucleic acid encodes a fusion polypeptide comprising a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof, and a polypeptide comprising at least 195 amino acids from the N-terminus of a MTB32A antigen (SEQ ID NO:2 or 4).

25. The expression cassette of claim 24, wherein the fusion polypeptide further comprises a polypeptide comprising at least 132 amino acids of the C-terminus of a MTB32A antigen (SEQ ID NO:2 or 4).

26. The expression cassette of claim 24, wherein the nucleic acid encodes a fusion polypeptide having the amino acid sequence of MTB59F (SEQ ID NO:20).

27. The expression cassette of claim 26, wherein the nucleic acid has the sequence of the nucleic acid encoding MTB59F (SEQ ID NO:19).

28. The expression cassette of claim 25, wherein the nucleic acid encodes a fusion polypeptide having the amino acid sequence of MTB72F (SEQ ID NO:16).

29. The expression cassette of claim 28, wherein the nucleic acid has the sequence of the nucleic acid encoding MTB72F (SEQ ID NO:15).

30. The expression cassette of claim 28, wherein the nucleic acid has the sequence of the nucleic acid encoding MTB72FMutSA (SEQ ID NO:18).

31. The expression cassette of claim 29or 30, further comprising a nucleic acid encoding at least one additional antigen from a *Mycobacterium* species of the tuberculosis complex, wherein the antigen is selected from the group consisting ofMTB8.4 antigen (SEQ ID NO:22), MTB9.8 antigen (SEQ ID NO:24), MTB9.9 antigen (SEQ ID NO:27), MTB40 antigen (SEQ ID NO:29), MTB41 antigen (SEQ ID NO:31), 38-1 (SEQ ID NO:35), TbRa3 (SEQ ID NO:37), 38 kD (SEQ ID NO:39), DPEP (SEQ ID NO:41), TbH4 (SEQ ID NO:43), DPPD(SEQ ID NO:45), MTB82, Erd14, ESAT-6 antigen (SEQ ID NO:33), MTB85 complex antigen, or α-crystalline antigen, or an immunogenic fragment thereof.

32. The expression cassette of claim 20, further comprising a nucleic acid encoding at least one additional antigen from a *Mycobacterium* species of the tuberculosis complex, wherein the antigen is selected from the group consisting ofMTB8.4 antigen (SEQ ID NO:22), MTB9.8 antigen (SEQ ID NO:24), MTB9.9 antigen (SEQ ID NO:27), MTB40 antigen (SEQ ID NO:29), MTB41 antigen (SEQ ID NO:31), 38-1 (SEQ ID NO:35), TbRa3 (SEQ ID NO:37), 38 kD (SEQ ID NO:39), DPEP (SEQ ID NO:41), TbH4 (SEQ ID NO:43), DPPD(SEQ ID NO:45), MTB82, Erd14, ESAT-6 antigen (SEQ ID NO:33), MTB85 complex antigen, or α-crystalline antigen, or an immunogenic fragment thereof.

33. The expression cassette of claim 20, further comprising a nucleic acid encoding an NS1 antigen.

34. The expression cassette of claim 20, wherein the *Mycobacterium* species is *Mycobacterium tuberculosis*.

35. A method for eliciting an immune response in a mammal, the method comprising the step of administering to the mammal an immunologically effective amount of a pharmaceutical composition comprising a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex, and a MTB32A antigen (SEQ ID NO:2 or 4) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex.

36. The method of claim 35, wherein the mammal has been immunized with BCG.

37. The method of claim 35, wherein the mammal is a human.

38. The method of claim 35, wherein the composition is administered prophylactically.

39. The method of claim 35, comprising a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex, and a polypeptide comprising at least 195 amino acids from the N-terminus of a MTB32A antigen (SEQ ID NO:2 or 4) from a *Mycobacterium* species of the tuberculosis complex.

40. The method of claim 39, further comprising a polypeptide comprising at least about 132 amino acids from the C-terminus of MTB32A antigen (SEQ ID NO: 2 or 4) from a *Mycobacterium* species of the tuberculosis complex.

41. The method of claim 35 or 39, wherein the antigens are covalently linked, thereby forming a fusion protein.

42. The method of claim 41, wherein the fusion polypeptide has the amino acid sequence of MTB59F (SEQ ID NO:20).

43. The method of claim 40, wherein the antigens are covalently linked, thereby forming a fusion protein.

44. The method of claim 43, wherein the fusion polypeptide has the amino acid sequence of MTB72F (SEQ ID NO:16).

45. The method of claim 43, wherein the fusion polypeptide has the amino acid sequence of MTB72FMutSA (SEQ ID NO:18).

46. The method of claim 35, wherein the pharmaceutical composition further comprises an adjuvant.

47. The method of claim 46, wherein the adjuvant comprises QS21 and MPL.

48. The method of claim 46, wherein the adjuvant is selected from the group consisting of AS2, ENHANZYN, MPL, 3D-MPL, IFA, QS21, CWS, TDM, AGP, CPG, Leif, saponin, and saponin mimetics.

49. A method for eliciting an immune response in a mammal, the method comprising the step of administering to the mammal an immunologically effective amount of an expression cassette comprising a nucleic acid encoding a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex, and a nucleic acid encoding a MTB32A antigen (SEQ ID NO:2 or 4) or an immunogenic fragment thereof from a *Mycobacterium* species of the tuberculosis complex.

50. The method of claim 49, wherein the mammal has been immunized with BCG.

51. The method of claim 49, wherein the mammal is a human.

52. The method of claim 49, wherein the composition is administered prophylactically.

53. The method of claim 49, wherein the nucleic acid encodes a fusion polypeptide comprising a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof, and a polypeptide comprising at least 195 amino acids from the N-terminus of a MTB32A antigen (SEQ ID NO:2 or 4) .

54. The method of claim 53, further comprising a nucleic acid encoding a polypeptide comprising at least 132 amino acids of the C-terminus of a MTB32A antigen (SEQ ID NO:2 or 4) from a *Mycobacterium* species of the tuberculosis complex.

55. The method of claim 49, wherein the nucleic acid encodes a fusion polypeptide comprising a MTB39 antigen (SEQ ID NO: 12 or 14) or an immunogenic fragment thereof and a nucleic acid encoding a MTB32A antigen (SEQ ID NO:2 or 4) or an immunogenic fragment thereof.

56. The method of claim 55, wherein the nucleic acid encodes a fusion polypeptide comprising a MTB39 antigen (SEQ ID NO:12 or 14) or an immunogenic fragment thereof, and a polypeptide comprising at least 195 amino acids from the N-terminus of a MTB32A antigen (SEQ ID NO: 2 or 4).

57. The method of claim 56, wherein the fusion polypeptide further comprises a polypeptide comprising at least 132 amino acids of the C-terminus of a MTB32A antigen (SEQ ID NO:2 or 4).

58. The method of claim 56, wherein the nucleic acid encodes a fusion polypeptide having the amino acid sequence of MTB59F (SEQ ID NO:20).

59. The method of claim 58, wherein the nucleic acid has the nucleotide sequence of the nucleic acid encoding MTB59F (SEQ IDNO:19).

60. The method of claim 57, wherein the nucleic acid encodes a fusion polypeptide having the amino acid sequence of MTB72F (SEQ ID NO:16) .

61. The method of claim 57, wherein the nucleic acid encodes a fusion polypeptide having the amino acid sequence of MTB72FMutSA (SEQ ID NO:18).

62. The method of claim 60, wherein the nucleic acid has the nucleotide sequence of the nucleic acid encoding MTB72F (SEQ IDNO:15).

63. The method of claim 60, wherein the nucleic acid has the nucleotide sequence of the nucleic acid encoding MTB72FMutSA (SEQ ID NO:17).

64. An isolated nucleic acid encoding a MTB32A antigen from a *Mycobacterium* species of the tuberculosis complex, wherein at least one amino acid in the active site triad of the MTB32A antigen (SEQ ID NO:2 or 4) has been substituted by a different amino acid.

65. The nucleic acid of claim 64, wherein an serine residue corresponding to amino acid position 183 of SEQ ID NO:4 or position 207 of SEQ ID NO:2 has been substituted by another amino acid.

66. The nucleic acid of claim 65, wherein an alanine residue has been substituted for the serine residue.

67. The nucleic acid of claim 66, wherein the nucleic acid comprises a nucleotide sequence of SEQ ID NO:5.

68. A composition comprising the nucleic acid of claim 64.

69. A nucleic acid encoding a fusion polypeptide comprising the nucleic acid of claim 64.

70. An isolated MTB32A polypeptide from a *Mycobacterium* species of the tuberculosis complex, wherein at least one amino acid in the active site triad of the MTB32A antigen (SEQ ID NO:2 or 4) has been substituted by a different amino acid.

71. The polypeptide of claim 70, wherein a serine residue corresponding to amino acid position 183 of SEQ ID NO:4 or amino acid position 207 of SEQ ID NO:2 has been substituted by another amino acid.

72. The polypeptide of claim 71, wherein an alanine residue has been substituted for the serine residue.

73. A polypeptide of claim 72, wherein the polypeptide comprises an amino acid sequence of SEQ ID NO:6.

74. A composition comprising the polypeptide of claim 70.

75. A fusion polypeptide comprising the polypeptide of claim 70.

76. An isolated nucleic acid encoding a fusion polypeptide comprising a MTB39 antigen (SEQ ID NO:12 or 14) from a *Mycobacterium* species of the tuberculosis complex, and an antigen comprising at least 195 amino acids from the N-terminus of a MTB32A antigen (SEQ ID NO:2 or 4) from a *Mycobacterium* species of the tuberculosis complex, wherein an amino acid of the active site triad of the MTB32A antigen (SEQ ID NO:2 or 4) has been substituted by a different amino acid.

77. The nucleic acid of claim 76, wherein a serine residue corresponding to amino acid at position 183 of SEQ ID NO:4 or position 207 or SEQ ID NO:2 has been substituted by another amino acid.

78. The nucleic acid of claim 77, wherein an alanine residue has been substituted for the serine residue.

79. A composition comprising the nucleic acid of claim 76.

80. A nucleic acid encoding a fusion polypeptide comprising the nucleic acid of claim 76.

81. A nucleic acid encoding a fusion polypeptide, wherein the nucleic acid comprises a nucleotide sequence of SEQ ID NO:17.

82. A nucleic acid encoding a fusion polypeptide comprising an amino acid sequence of SEQ ID NO:18.

83. An isolated polypeptide encoding a fusion polypeptide comprising a MTB39 (SEQ ID NO: 12 or 14) antigen from a *Mycobacterium* species of the tuberculosis complex, and an antigen comprising at least 195 amino acids from the N-terminus of a MTB32A antigen (SEQ ID NO:2 or 4) from a *Mycobacterium* species of the tuberculosis complex, wherein an amino acid of the active site triad of the MTB32A antigen (SEQ ID NO:2 or 4) has been substituted by a different amino acid.

84. The polypeptide of claim 83, wherein an serine residue corresponding to amino acid position 183 of SEQ ID NO:4 or amino acid position 207 of SEQ ID NO:2 has been substituted by another amino acid.

85. The polypeptide of claim 83, wherein an alanine residue has been substituted for the serine residue.

86. A composition comprising the polypeptide of claim 83.

87. A fusion polypeptide comprising the polypeptide of claim 83.

88. A fusion polypeptide comprising an amino acid sequence of SEQ ID NO:18.

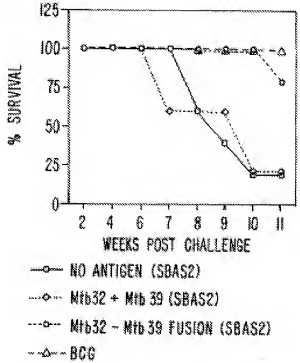

125
100
75
50
25
0
% SURVIVAL
2
4
6
7
8
9
10
11
WEEKS POST CHALLENGE
NO ANTIGEN (SBAS2)
Mtb32 + Mtb 39 (SBAS2)
Mtb32 - Mtb 39 FUSION (SBAS2)
BCG

FIG. 1.

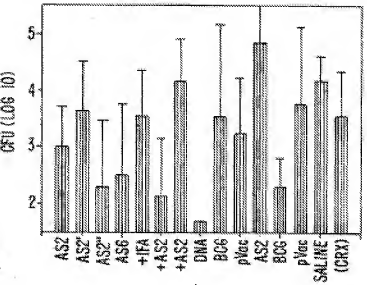

CFU (LOG 10)
5
4
3
2
AS2
AS2'
AS2"
AS6
+IFA
+AS2
+AS2
DNA
BCG
pVac
AS2
BCG
pVac
SALINE
(ORX)

FIG. 2A.

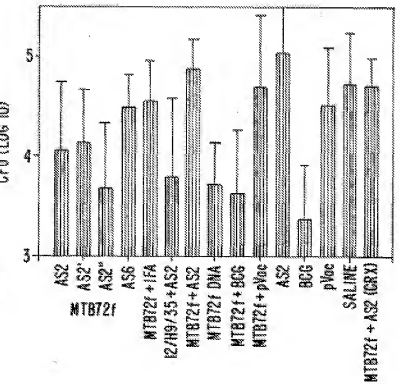

CFU (LOG 10)
5
4
3
AS2
AS2'
AS2"
AS6
MTB72f +IFA
12/H9/35+AS2
MTB72f+AS2
MTB72f DNA
MTB72f+BCG
MTB72f+pVac
AS2
BCG
pVac
SALINE
MTB72f + AS2 (ORX)
MTB72f

FIG. 2B.

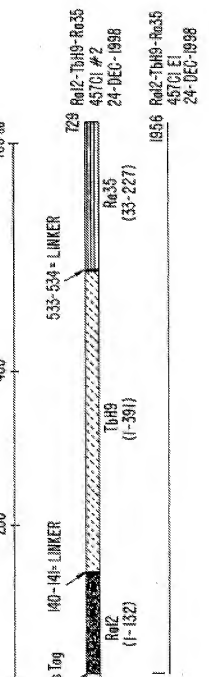

200
400
700 aa
His Tag
140-141= LINKER
533-534= LINKER
T29 Ra12-TbH9-Ra35
457C1 #2
24-DEC-1998
Ra12
(1-132)
TbH9
(1-391)
Ra35
(33-227)
1956 Ra12-TbH9-Ra35
457C1 E1
24-DEC-1998

FIG. 3.

Ra35 N-terminus DNA

```
gcccgccgg cctgtgcga ggaccggttc gccgactccc ccgcgctgcc cctcgacccg tccgcgatgg 70
tcgcccaagt gggccacag gtggtcaaca tcaacaccaa actgggctac aacaagcccg tgggcgccgg 140
gacgggcatc gtcatcgatc ccaccggtgt cgtgctgacc aacaccacg tgatcgtggg cgccaccgac 210
atcaatgcgt tcagcgtcgg ctccggccaa acctacggcg tcgatgtggt cgggtatgac cgcaccgagg 280
atgtcgcggg gctgcagctg cgcgtgccg gtggcctacc atcgccgggc atcggtggcg gcgtcgcggt 350
tggtgagccc gtcgtcgcga tggccaacag cggtgggcag ggcgtgaccg cccgtgcgct gcctggcagg 420
gtggtcgcgc tcggccaaac cgtgcaggcg tcggatcggc tgaccggtgc cgaagagaca ttgaacggct 490
tgatccagtt cgatgccgcg atccagcccg gtgatcgggg cgggcccgtc gtcaacggcc tagyacaggt 560
ggtcggtatg aacacggccg cgtcctag 588
```

Ra35 N-terminus amino acid sequence

```
Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala
 5                  10                  15                  20
Met Val Ala Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val
            25                  30                  35                  40
Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val Ile Ala
        45                  50                  55                  60
Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly
    65                  70                  75                  80
Tyr Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala
                85                  90                  95                  100
                                                    105                 110
```

FIG. 4.

```
Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly Gly Gln Gly Gly
                115                 120                 125                 130

Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu
        135                 140                 145                 150

Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
155                 160                 165                 170                 175

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr Ala Ala Ser
                180                 185                 190                 195
```

FIG. 4. (CONTINUED)

```
         Ra12
1    MHHHHHHTAASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLG Mtb72f
1    MHHHHHHTAASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLG Mtb72f-mutSA

56   LGVVDNNGNGARVQRVVGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHH Mtb72f
56   LGVVDNNGNGARVQRVVGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHH Mtb72f-mutSA
                                             TbH9FL
111  PGDVISVTWQTKSFFTRTFNVTLAEGPPAEFMVDFGALPPEINSARMYAGPGSAS Mtb72f
111  PGDVISVTWQTKSFFTRTFNVTLAEGPPAEFMVDFGALPPEINSARMYAGPGSAS Mtb72f-mutSA

166  LVAAAQMWDSVASDLFSAASAFQSVVWGLTVGSWIGSSAGLMVAAASPYVAWMSV Mtb72f
166  LVAAAQMWDSVASDLFSAASAFQSVVWGLTVGSWIGSSAGLMVAAASPYVAWMSV Mtb72f-mutSA

221  TAGQAELTAAQVRVAAAAYETAYGLTVPPPVIAENRAELMILIATNLLGQNTPAI Mtb72f
221  TAGQAELTAAQVRVAAAAYETAYGLTVPPPVIAENRAELMILIATNLLGQNTPAI Mtb72f-mutSA

276  AVNEAEYGEMWAQDAAAMFGYAAATATATATLLPFEEAPEMTSAGGLLEQAAAVE Mtb72f
276  AVNEAEYGEMWAQDAAAMFGYAAATATATATLLPFEEAPEMTSAGGLLEQAAAVE Mtb72f-mutSA

331  EASDTAAANQLMNNVPQALQQLAQPTQGTTPSSKLGGLWKTVSPHRSPISNMVSM Mtb72f
331  EASDTAAANQLMNNVPQALQQLAQPTQGTTPSSKLGGLWKTVSPHRSPISNMVSM Mtb72f-mutSA

386  ANNHMSMTNSGVSMTNTLSSMLKGFAPAAAAQAVQTAAQNGVRAMSSLGSSLGSS Mtb72f
386  ANNHMSMTNSGVSMTNTLSSMLKGFAPAAAAQAVQTAAQNGVRAMSSLGSSLGSS Mtb72f-mutSA

441  GLGGGVAANLGRAASVGSLSVPQAWAAANQAVTPAARALPLTSLTSAAERGPGQM Mtb72f
441  GLGGGVAANLGRAASVGSLSVPQAWAAANQAVTPAARALPLTSLTSAAERGPGQM Mtb72f-mutSA
```

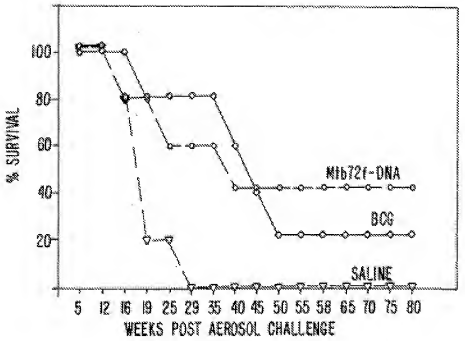
FIG. 5.

```
                                                                  Ra35
496 LGGLPVGQMGARAGGGLSGVLRVPPRPYVMPHSPAAGD[HAPPALSQDRFADFPAL Mtb72f
496 LGGLPVGQMGARAGGGLSGVLRVPPRPYVMPHSPAAGD[HAPPALSQDRFADFPAL Mtb72f-mutSA

551 PLDPSAMVAQVGPQVVNINTKLGYNWAVGAGTGIVIDPNGVVLTNNHVIAGATDI Mtb72f
551 PLDPSAMVAQVGPQVVNINTKLGYNNAVGAGTGIVIDPNGVVLTNNHVIAGATDI Mtb72f-mutSA

606 NAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGN Mtb72f
606 NAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGN Mtb72f-mutSA

661 SGGQGGTPRAVPGKVVALGQTVQASDSLNGAEETLNGLIQFDAAIQPGDSGGPVV Mtb72f
661 SGGQGGTPRAVPGKVVALGQTVQASDSLTGAEETLNGLIQFDAAIQPGDSGGPVV Mtb72f-mutSA

716 NGLGQVVGMNTAAS                                          Mtb72f
716 NGLGQVVGMNTAAS                                          Mtb72f-mutSA
```

FIG. 5. (CONTINUED)

```
                    Ra35 N-term
1   MHHHHHHMPPALSQDRFADFPALPLDPSAMVAQVGPQVVNINTKLGYNNA TbRa35_mat
1   MHHHHHHMPPALSQDRFADFPALPLDPSAMVAQVGPQVVNINTKLGYNNA TbRa35_mutSA

51  VGAGTGIVIDPNGVVLTNNHVIAGATDINAFSVGSGQTVGVDVVGYDRTQ TbRa35_mat
51  VGAGTGIVIDPNGVVLTNNHVIAGATDINAFSVGSGQTVGVDVVGYDRTQ TbRa35_mutSA

101 DVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGNSGGQGGTPRAVPGRVVA TbRa35_mat
101 DVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGNSGGQGGTPRAVPGRVVA TbRa35_mutSA
                                                  Ra12 Cterm
151 LGQTVQASDSITGAEETLNGLIQFDAAIQPGDSGGPVVNGLGQVVGMNTA TbRa35_mat
151 LGQTVQASDSITGAEETLNGLIQFDAAIQPGDSGGPVVNGLGQVVGMNTA TbRa35_mutSA
      end Ra35 Nterm
201 ASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLGLGVV TbRa35_mat
201 ASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLGLGVV TbRa35_mutSA

251 DNNGNGARVQRVVGSAPAASIGIGTGDVITAVDGAPINSATAMADALNGH TbRa35_mat
251 DNNGNGARVQRVVGSAPAASIGIGTGDVITAVDGAPINSATAMADALNGH TbRa35_mutSA

301 HPGDVISVTWQTKSGGTRTGNVTLAEGPPA end                  TbRa35_mat
301 HPGDVISVTWQTKSGGTRTGNVTLAEGPPA Ra12                 TbRa35_mutSA
```

FIG. 6.


% SURVIVAL
100
80
60
40
20
Mtb72f-DNA
BCG
SALINE
5 12 16 19 25 29 35 40 45 50 55 58 65 70 75 80
WEEKS POST AEROSOL CHALLENGE

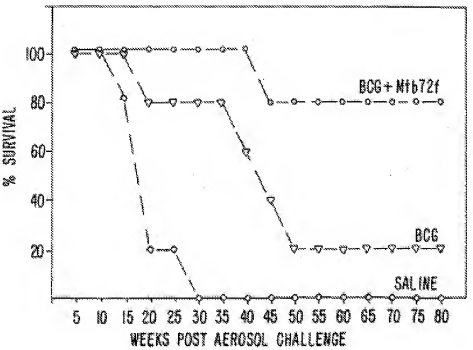

% SURVIVAL
100
80
60
40
20
BCG+Mtb72f
BCG
SALINE
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
WEEKS POST AEROSOL CHALLENGE

FIG. 7

SEQUENCE LISTING

(2) INFORMATION FOR SEQ ID NO:1: MTB32A (Ra35 FL)

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1872 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GACTACGTTG GTGTAGAAAA ATCCTGCCGC CCGGACCCTT AAGGCTGGGA CAATTTCTGA     60
TAGCTACCCC GACACAGGAG GTTACGGGAT GAGCAATTCG CGCCGCCGCT CACTCAGGTG    120
GTCATGGTTG CTGAGCGTGC TGGCTGCCGT CGGGCTGGGC CTGGCCACGG CGCCGGCCCA    180
GGCGGCCCCG CCGCCCTTGT CGCAGGACCG GTTCGCCGAC TTCCCCGCGC TGCCCCTCGA    240
CCCGTCCGCG ATGGTCGCCC AAGTGGGGCC ACAGGTGGTC AACATCAACA CCAAACTGGG    300
CTACAACAAC GCCGTGGGCG CCGGGACCGG CATCGTCATC GATCCCAACG GTGTCGTGCT    360
GACCAACAAC CACGTGATCG CGGGCGCCAC CGACATCAAT GCGTTCAGCG TCGGCTCCGG    420
CCAAACCTAC GGCGTCGATG TGGTCGGGTA TGACCGCACC CAGGATGTCG CGGTGCTGCA    480
GCTGCGCGGT GCCGGTGGCC TGCCGTCGGC GGCGATCGGT GGCGGCGTCG CGGTTGGTGA    540
GCCCGTCGTC GCGATGGGCA ACAGCGGTGG GCAGGGCGGA ACGCCCCGTG CGGTGCCTGG    600
CAGGGTGGTC GCGCTCGGCC AAACCGTGCA GGCGTCGGAT TCGCTGACCG GTGCCGAAGA    660
GACATTGAAC GGGTTGATCC AGTTCGATGC CGCAATCCAG CCCGGTGATT CGGGCGGGCC    720
CGTCGTCAAC GGCCTAGGAC AGGTGGTCGG TATGAACACG GCCGCGTCCG ATAACTTCCA    780
GCTGTCCCAG GGTGGGCAGG GATTCGCCAT TCCGATCGGG CAGGCGATGG CGATCGCGGG    840
CCAAATCCGA TCGGGTGGGG GGTCACCCAC CGTTCATATC GGGCCTACCG CCTTCCTCGG    900
CTTGGGTGTT GTCGACAACA ACGGCAACGG CGCACGAGTC CAACGCGTGG TCGGGAGCGC    960
TCCGGCGGCA AGTCTCGGCA TCTCCACCGG CGACGTGATC ACCGCGGTCG ACGGCGCTCC   1020
GATCAACTCG GCCACCGCGA TGGCGGACGC GCTTAACGGG CATCATCCCG GTGACGTCAT   1080
CTCGGTGAAC TGGCAAACCA AGTCGGGCGG CACGCGTACA GGGAACGTGA CATTGGCCGA   1140
GGGACCCCCG GCCTGATTTG TCGCGGATAC CACCGGCCGG CCGGCCAATT GGATTGGCGC   1200
CAGCCGTGAT TGCCGGCGGA GCCCCCGAGT TCCGTCTTCC GTGCGCGGGG CGTTGTGGAA   1260
GCAATGAACG AGGCAGAACA CAGCGTTGAG CACCCTCCCG TGCAGGGCAG TTACGTCGAA   1320
GACGGTGTGG TCGAGCATCC GGATGCCAAG GACTTCGGCA GCGCGCCCGC CCTGCCCGCC   1380
GATCCGACCT GGTTTAAGCA CGCCGTCTTC TACGAGGTGC TGGTCCGGGC GTTCTTCGAC   1440
GCCAGCGCGG ACGGTTCCGG CGATCTGCGT GGACTCATCG ATCGCCTCGA CTACCTGCAG   1500
TGGCTTGGCA TCGACTGCAT CTGTTTGCCG CGTTCCTACG ACTCACCGCT GCGCGACGGC   1560
GGTTACGACA TTCGCGACTT CTACAAGGTG CTGCCCGAAT TCGGCACCGT CGACGATTTC   1620
GTCGCCCTGG TCGACACCGC TCACCGGCGA GGTATCCGCA TCATCACCGA CCTGGTGATG   1680
AATCACACCT CGGAGTCGCA CCCCTGGTTT CAGGAGTCCC GCCGCGACCC AGACGGACCG   1740
TACGGTGACT ATTACGTGTG GAGCGACACC AGCGAGCGCT ACACCGACGC CCGGATCATC   1800
TTCGTCGACA CCGAAGAGTC GAACTGGTCA TTCGATCCTG TCCGCCGACA GTTCTACTGG   1860
GCACCGATTC TT                                                       1872
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 355 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2: MTB32A (Ra35FL)

```
Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
1               5                   10                  15
Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
            20                  25                  30
Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
        35                  40                  45
Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val
    50                  55                  60
Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
65                  70                  75                  80
```

```
Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
                85                  90                  95
Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
            100                 105                 110
Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
        115                 120                 125
Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
    130                 135                 140
Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
145                 150                 155                 160
Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
                165                 170                 175
Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
            180                 185                 190
Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
        195                 200                 205
Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
    210                 215                 220
Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala
225                 230                 235                 240
Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly
                245                 250                 255
Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu
            260                 265                 270
Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val
        275                 280                 285
Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile
    290                 295                 300
Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp
305                 310                 315                 320
Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp Gln
                325                 330                 335
Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
            340                 345                 350
Pro Pro Ala
        355
```

<212> DNA
<213> Ra35 mature
<400> SEQ ID NO:3

```
catatgcatc accatcacca tcacggcccg ccggccttgt cgcaggaccg gttcgccgac 60
ttccccgcgc tgcccctcga cccgtacgcg atggtcgccc aagtgggccc acaggtggtc 120
aacatcaaca cccaactggg ctacaacaac gccgtgggcg ccgggaccgg catcgtcatc 180
gatcccaacg gtgtcgtgct gaccaacaac cacgtgatcg cgggcgccac cgacatcaat 240
gcgttcagcg tcggctccgg ccaaacctac ggcgtcgatg tggtcgggta tgaccgcacc 300
caggatgtcg cggtgctgca gctgcgcggt gccggtggcc tgccgtcggc ggcgatcggt 360
ggcggcgtcg cggttggtga gcccgtcgtc gcgatgggca acagcggtgg gcagggcgga 420
acgccccgtg cggtgcctgg cagggtggtc gcgctcggcc aaaccgtgca ggcgtcggat 480
tcgctgaccg gtgccgaaga gacattgaac gggttgatcc agttcgatgc cgcgatccag 540
cccggtgacg cgggcgggcc cgtcgtcaac ggcctaggac aggtggtcgg tatgaacacg 600
gccgcgtccg ataacttcca gctgtcccag ggtgggcagg gattcgccat tccgatcggg 660
caggcgatgg cgatcgcggg ccagatccga tcgggtgggg ggtcacccac cgttcatatc 720
ggccctaccg ccttcctcgg cttgggtgtt gtcgacaaca acggcaacgg cgcacgagtc 780
caacgcgtgg tcggcagcgc tccggcggca agtctcggca tctccaccgg cgacgtgatc 840
accgcggtcg acggcgctcc gatcaacctg gccaccgcga tggcggacgc gcttaacggg 900
catcatcccg gtgacgtcat ctcggtgacc tggcaaacca agtcgggcgg cacgcgtaca 960
gggaacgtga cattggccga gggacccccg gcctgagaat tc                    1002
```

<212> PRT
<213> Ra35 mature
<400> SEQ ID NO:4

```
Met His His His His His His Ala Pro Pro Ala Leu Ser Gln Asp Arg
                  5                  10                  15
```

```
Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala
            20                  25                  30

Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn
        35                  40                  45

Asn Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val
    50                  55                  60

Val Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala
65                  70                  75                  80

Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr
                85                  90                  95

Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly
            100                 105                 110

Leu Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val
        115                 120                 125

Val Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val
    130                 135                 140

Pro Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser
145                 150                 155                 160

Leu Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala
                165                 170                 175

Ala Ile Gln Pro Gly Asp Ser Gly Gly Pro Val Val Asn Gly Leu Gly
            180                 185                 190

Gln Val Val Gly Met Asn Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser
        195                 200                 205

Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile
    210                 215                 220

Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly
225                 230                 235                 240

Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly
                245                 250                 255

Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly
            260                 265                 270

Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn
        275                 280                 285

Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp
    290                 295                 300

Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly
305                 310                 315                 320

Asn Val Thr Leu Ala Glu Gly Pro Pro Ala
                325                 330
```

<212> DNA
<213> Ra35FLMutSA
<400> SEQ ID NO:5

```
catatgcatc accatcacca tcacgccccg ccggccttgt cgcaggaccg gttcgccgac 60
ttcccggcgc tgcccctcga cccgtccgcg atggtcgccc aagtcggccc ccaggtggtc 120
aacatcaaca ccaaactggg ctacaacaac gccgtcggcg ccggcaccgg catcgtcatc 180
```

```
gatcccaacg gtgtcgtgct gaccaacaac cacgtgatcg cgggcgccac cgacatcaat  240
gcgttcagcg tcggctccgg ccaaacctac ggcgtcgatg tcgtcgggta tgaccgcacc  300
caggatgtcg cggtgctgca gctgcgcggt gccggtggcc tgccgtcggc ggcgatcggt  360
ggcggcgtcg cggttggtga gcccgtcgtc gcgatgggca acagcggtgg gcagggcgga  420
acgccccgtg cggtgcctgg cagggtggtc gcgctcggcc aaaccgtgca ggcgtcggat  480
tcgctgaccg gtgccgaaga gacattgaac gggttgatcc agttcgatgc cgcgatccag  540
cccggtgatt cgggcgggcc cgtcgtcaac ggcctaggac aggtggtcgg tatgaacacg  600
gccgcgtccg ataacttcca gctgtcccag ggtgggcagg gattcgccat tccgatcggg  660
caggcgatgg cgatcgcggg ccagatccga tcgggtgggg ggtcacccac cgttcatatc  720
gggcctaccg ccttcctcgg cttgggtgtt gtcgacaaca acggcaacgg cgcacgagtc  780
caacgcgtgg tcgggagcgc tccggcggca agtctcggca tctccaccgg cgacgtgatc  840
accgcggtcg acggcgctcc gatcaactcg gccaccgcga tggcggacgc gcttaacggg  900
catcatcccg gtgacgtcat ctcggtgacc tggcaaacca agtcgggcgg cacgcgtaca  960
gggaacgtga cattggccga gggacccccg gcctgagaat tc                     1002
```

<212> PRT
<213> Ra35FLMutSA
<400> SEQ ID NO:6

```
Met His His His His His His Ala Pro Pro Ala Leu Ser Gln Asp Arg
                    5                  10                  15

Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala
            20                  25                  30

Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn
        35                  40                  45

Asn Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val
    50                  55                  60

Val Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala
65                  70                  75                  80

Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr
                85                  90                  95

Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly
            100                 105                 110

Leu Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val
        115                 120                 125

Val Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val
    130                 135                 140

Pro Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser
145                 150                 155                 160

Leu Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala
                165                 170                 175

Ala Ile Gln Pro Gly Asp Ala Gly Gly Pro Val Val Asn Gly Leu Gly
            180                 185                 190

Gln Val Val Gly Met Asn Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser
        195                 200                 205

Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile
    210                 215                 220

Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly
225                 230                 235                 240

Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly
                245                 250                 255
```

Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly
260 265 270

Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn
275 280 285

Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp
290 295 300

Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly
305 310 315 320

Asn Val Thr Leu Ala Glu Gly Pro Pro Ala
325 330

(2) INFORMATION FOR SEQ ID NO:7: Ra35 (MTB32A N-term)

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 615 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

gccccgccggccttgtcgcaggaccggttcgccgacttccccgcgctgcccctcgacccgtccgcg
atggtcgcccaagtggcgccgcaggtggtcaacatcaacaccaaactgggctacaacaacgccgtg
ggcgccgggaccggcatcgtcatcgatcccaacggtgtcgtgctgaccaacaaccacgtgatcgcg
ggcgccaccgacatcaacgcgttcagcgtcggctccggccaaacctacggcgtcgatgtggtcggg
tatgaccgcacccaggatgtcgcggtgctgcagctgcgcggtgccggtggcctgccgtcggcggcg
atcggtggcggcgtcgcggttggtgagcccgtcgtcgcgatgggcaacagcggtgggcagggcgga
acgccccgtgcggtgcctggcagggtggtcgcgctcggccaaaccgtgcaggcgtcggattcgctg
accggtgccgaagagacattgaacgggttgatccagttcgatgccgcgatccagcccggtgattcg
ggcgggcccgtcgtcaacggcctaggacaggtggtcggtatgaacacggccgcgtcc

(2) INFORMATION FOR SEQ ID NO:8: Ra35 (MTB32A N-term)

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 205 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val

Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr

Ala Ala Ser

(2) INFORMATION FOR SEQ ID NO:9: Rat2

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 447 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGGTATGAAC AGGCGTCGTA CGAATAACTT CCAGCTTTCC CAGGGTGGCC AGGAATTCGC 60
CATTCGGATC GGGCAGGCGA TGGCGATCGC GGGCGAGATC CGATCGGTTG GGTGGTCAGC 120
CAGCTTCAT ATCGGGCCTA CGCCTTCCT CCTCTTGCT GTTGTTGACA ACAACGGCAA 180
CGGCGCACGA GTCCAAACGC TGGTCGGAAG CGCTTCGCGA GCAAGTCTCG GCATCTCCAC 240
CGGCAACGTG ATCACCGCGG TCAACGGCGC TCGGATCAAC TCGGCCACCG CGATCGCGGA 300
CGCGCTTAAC GGTCATCATC CGCGTGACAT CATCTCGGTG AACTGGCAAA CCAAGTCGAG 360
CGGCAGCGGT ACAAGCAACG TGACATTTAC CAAGGACCC CCGGCTTCAT TTCGTCGTGC 420
ATAGCACCCG CCGCCGCGCC AATTGGA 447

(2) INFORMATION FOR SEQ ID NO:10: Rat2

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 132 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
1 5 10 15
Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
20 25 30
Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
35 40 45
Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
50 55 60
Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
65 70 75 80
Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
85 90 95
Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp
100 105 110
Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
115 120 125
Gly Pro Pro Ala
130

(2) INFORMATION FOR SEQ ID NO:11: TBP3

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 851 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTGCAGGGTG GCCCCGATCA GCCTCAGCGC CGGCCAGGCC GAGCTGACCG CCGCCCAGGT 60
CGGGTTGGT GCGGCGGCCT ACGATACGGC GTATGGCTG ACGGTGCCCC CGCCGGTAT 120

```
CGCCGAGAAC CATGCTGAAC TGATGATTCT GATAGCGACC AACTTCTTGG GGCAAAACAC     180
CCGGCCGATC GCGGTCAACG AGGCCGATA CGGCGAGGTG TGGCCCCAAG ACGCCGCGGC     240
GATGTTTGGC TACGCCGGGT CGACGGCGAC GGCGACGGCG ACGTTGCTGC CGTTCGAGGA     300
GGCGCCGAAG ATGACGACGA CGGGTGGGCC GCTCGGCAAG GCCGCAGCGG TCGAGGAGCC     360
CTCCGACACC GCCGCGGCGA ACCAGTTGAC GRACAATGTG CGCCAGGGCC TGAAACAGTT     420
GGCCCAGCCC ACGCAGGGCA CCACGCCTTC CTCCAAGCTG GGTTGCCTGT GCAAGACGGT     480
CTCGCCGCAT GGGTCGCCGA TCAGCAAGAT GGTGTCGATG GCCAACAACC ACATYTCCAT     540
GACCAACTCG GGTGTTGTCG TGACCAACAC CTTCGACTCG ATGCTGAAGG GCTTCTGCTC     600
GGCGGCTGGC GCCGCAGCCG TGCAAACCGC GCGCCAAAAC GGGGTCCTGG CGATGAGCTC     660
GCTGGGCAGC TCGCTGGCGT CTTCGGTCT GTCCGGTCGG GTCGCCGCCA ACTTGGGTCG     720
GGCGGCCTCG GTACGGTATG GTCACCGGCA TGGCGGAAAA TATGCAAAGT CTGGTCGGCG     780
GRACGGTTGT CGGCGTAAG GTTTACCGCC GTTTTCTGGA TGCGGTAAAC TTCGTTCAACG     840
GRAACAGTTA C                                                          851
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 263 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12: TRP9

```
Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
1               5                  10                  15
Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
            20                  25                  30
Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
        35                  40                  45
Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
    50                  55                  60
Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
65                  70                  75                  80
Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
                85                  90                  95
Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
            100                 105                 110
Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
        115                 120                 125
Asn Asn Val Pro Gln Ala Leu Lys Gln Leu Ala Glu Pro Thr Glu Gly
    130                 135                 140
Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
145                 150                 155                 160
His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
                165                 170                 175
Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
            180                 185                 190
Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
        195                 200                 205
Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
    210                 215                 220
Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
225                 230                 235                 240
Ser Val Arg Tyr Gly His Arg Asp Gly Gly Lys Tyr Ala Xaa Ser Gly
                245                 250                 255
Arg Arg Asn Gly Gly Pro Ala
            260
```

(2) INFORMATION FOR SEQ ID NO:13: TRHDPL

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3058 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GATCGTACCC GTGCCAGTGC TCGGGCCGTT TGAGGATTGA GTGCACGTGT CTTTCGTGAT 60

GGCATACCCA GAGATGTTGG CGGCGGCGGC TGACACCCTG CAGAGCATCG GTGCTACCAC 120

TGTGGCTAGC AATGCCGCTG CGGCGGCTCC GACGACTGGG GTGTTGGCCC CCGCTGCGGA 180

TGAGGTGTCG GCGCTGACTG CGGCGACTT CGCCGCAGAT GCGGCGATGT ATCAGTCCGT 240

GAGCAGTAGG GCTGCTGCGA TTCATGACCA GTTCGTGGCC ACTCTTGCCA GCAGCGCCAG 300

CTGGTATGCG GCTACTGAAG TGGCTAATGC GGCGGCGGCC AGCTAAGCCA GGAACAGTCG 360

GCACGAGAAA CCACGAGAAA TAGGGACACC TAATGGTTGA TTTCGGGGCG TTACCACCGG 420

AGATCAACTC GCCGAGGATG TACGCGGCC CTGGTTCGGC CTGCTCTGTG GCGCGGTCC 480

AGATGTCGGA CAGCGTGGCG AGTCACCTGT TTTCGCCCGC GTCGCGTTTT CAGTCGGTGG 540

TCTGGGGTCT GACGGTGGGG TCGTCGATAG GTTCGTCGGC GGGTCTGATG GTGGCGGCGG 600

CGTCGCGCTA TATGGCGTGG ATGAGCGTTA CCGCGGGGCA GACCGAGCTG ACGCCGCCC 660

AGTCCGGGT TGCTCGCGCG GCCTACGAGA CGGCGATATG GCTGACGGTG CCCCGGCGG 720

TGATCGCGA GAACCGTGCT GAACTGATGA TTCTGATAGC GACCAACCC TTGGGGCAAA 780

ACACCCCGCC GATCGCGTC AACGAGGCCG AATACGCGGA GATTTGGAGC CAAGACGCG 840

CGGGATGTT TGCCTACGCC GCGCGAGAGG CGACGACGAC GTCGACGTTG CTGCGTTCG 900

AGGAGGCGCC GGAGATGACC AGCGCGGTG GCCTCCTCGA GCAGGCGCC GCGGTCGAGG 960

AGGCTTCGA CACCGCCGCG GCGAACGAGT TGATGAACAA TGTGCCCCAG GCGCTGCAAC 1020

AGCTGGCCA GCCCAGCGAG GCCACCACGC CTTCTTCTAA GCTGGGTGCC CTGTGGAAGA 1080

CGGTCTGCC GCATCGGTCG CCGATCAGCA ACATGGTGTC GATGGCCAAC AAGCAGATGT 1140

CGATGACCAA CTCGGTGTTC TCGATGACCA ACACCTTGAG CTCGATGTTG AAGGCTTTG 1200

CTCCGGCGC GCCGGCCAG GCGGTGCAAA CCGCGGCGCA AAACGGGTC CGGGCGATGA 1260

GCTCGCTGGG CAGCTCGCTG GGTTCTTCGG GTCTGGCGG TGGGTGGCC GCCAACTTCG 1320

TTCGGCGCG CTCGGTCGGT TCGTTGTCGG TGCCGCAGGC CTGGCCGCG GCCAACCAG3 1380

CAGTCACCCC GGCGGCGCGG GCGCTGCCGC TGACCAGCCT GACCAGCGCC GCGGAAAGAG 1440

GGCCGGGCA GATGCTGGGC GGCCTGCTGG TGGGCAGAT GGGCGCCAGG GCGGTTGTTG 1500

GGCTCAGTGG TGTGCTGCGT GTTCGCCCGC GACCCTAIGT GATGCCGCAT TCTCCGGCGG 1560

CCGGCTAGGA GACGGGCGC AGACTGTCGT TATTTGACGA GTCATCGGCG GTCTCGGTGT 1620

TTCCGCGGCC GGCTATGACG ACAGTCAATG TGCATGACAA GTTACAGGTA TTAGGTCCAG 1680

GTTCAACAAG GAGACAGGCA ACATTGGCCC ACGTTTATG ACGCATCCGC AGCGATGCG 1740

GGACATGCG GGCGTTTTTG AGGTGCAGC CCAGACGTG GAGGACGAGG CTGGCCGGAT 1800

GTCGCGCTCC GCGCAAAACA TTTCCGGTGC GGACTGGAGT GGCATGGCCG AGCGWACCTC 1860

GCTAGACAGC ATGGCCCAGA TGAATCAGGC GTTTCGCAAC ATCGTGAACA TGCTCCAGGG 1920

GGTGCGTGAC GGGCTGGTTC GCGACGCCAA CAACTACGAG CAGCAGAGAC AGGCTCCCCA 1980

GCAGATCCTC AGCAGCIAAC GTCAGCCGCT GCAGCACAAT ACTTTTACAA GCGAAGGAGA 2040

ACAGTTCGA TGACCATCAA CTATCAATTC GGCGATGTCG ACACTCAGGG CGCCATAGTC 2100

CGCACTCAAG CCGCGTTTCT CGAAGCCGAG CATCAGGCCA TCATTCGTGA TGTGTTGACC 2160

GCGAGTGACT TTTCGGGCGG CGCCGGTTCG GCGCCTGCC ACGGGTTCAT TACCCAGTTG 2220

GCCCGTAACT TCCAGGTCAT CTACGAGCAG GCCAAGCCC ACGGCCAGAA GGTCGAGCT 2280

GCCGGCAACA ACATCGCCCA AACCGACAGC GCCGTCGGCT CCAGCTCGGC CTGACACCAG 2340

GCCAAGCCA GGAAGGTGGT GTACGAGTGA AGTTCCTCGC GTGATCCTTC GGGTGGCAGT 2400

CTAAGTGGTC AGTGCTGGCG TGTTGTGAGT TTGCTGCTTG GCGGGTTCTT CAGTGCTGGT 2460

CAGTGCTGCT CAGGCTCGGG TGAGGACCTC GAGGCCCAGG TAGCGCGTTC CTTCGAATCA 2520

TTCGTCGTGT TGTTCGGCGA GTACGGCTCC GACGAGGCGG ATGATCAAGG CGCGTTCGGG 2580

GAAGATGCGC ACGACGTCGG TTCGGCGTCG TACCTCTCGG TTGAGGCGTT CCTGGCGGTT 2640

GTTGTAGCAG ATTTGGCGCC AGATCTGCTT GCGGAAGGCG GTGAACGCCA GCAGGTCGGT 2700

GCGGCGGGTG TCGAGGTGCT CGGCCACCGC GGGAAGTTTG TCGGTCAGAG CGTCGAATAC 2760

CCGATCATAT TGAGCAACAA CTGATTCAGC GTCGGACTGG TCGTAGATGG AGTGCAGCAG 2820

GTTGCGCACC CACGGCCAGG AGGACTTCGG GGTGGCTGCC ATCAGATTGG CTGCGTAGTG 2880

GGTTCGCAG CGTTGCCAGG CGGCTGGGGG CAGGGTGCGC CGGATCGCGG CCACCAGGCC 2940

GGCGTGGGCG TCGCTGGTGA CCAGCGCGAC CCCGGACAGG CGCCGGCGCA CCAGGTCGCG 3000

GAAGAACGCC AGCCAGTCGG CCCGTCCTTC GGCGGAGTTG ACCTGGATGC CCAGGATC 3058

(2) INFORMATION FOR SEQ ID NO:14: THERE

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 191 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
1 5 10 15

Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp
20 25 30

Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
35 40 45

Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
50 55 60

Leu Met Val Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
65 70 75 80

Ala Gly Gln Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala
85 90 95

Ala Tyr Glu Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala
100 105 110

```
Glu Asn Arg Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly
            115                 120                 125

Gln Asn Thr Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Gln Met
        130                 135                 140

Trp Ala Gln Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala
145                 150                 155                 160

Thr Ala Thr Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr
                165                 170                 175

Ser Ala Gly Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser
            180                 185                 190

Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu
        195                 200                 205

Gln Gln Leu Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu
    210                 215                 220

Gly Gly Leu Trp Lys Thr Val Ser Pro His Arg Ser Pro Ile Ser Asn
225                 230                 235                 240

Met Val Ser Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val
                245                 250                 255

Ser Met Thr Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala
            260                 265                 270

Ala Ala Ala Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala
        275                 280                 285

Met Ser Ser Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly
    290                 295                 300

Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val
305                 310                 315                 320

Pro Gln Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg
                325                 330                 335

Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly
            340                 345                 350

Gln Met Leu Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly
        355                 360                 365

Gly Gly Leu Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met
    370                 375                 380

Pro His Ser Pro Ala Ala Gly
385                 390
```

<210> SEQ ID NO:15
<211> 2287
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence:tri-fusion protein Mtb72F(Ra12-TbH9-Ra35 or Mtb32-Mtb39 fusion)

```
tctagaaata attttgttta ctttaagaag gagatatacc t atg cat cac cat cac cac 56
                                                Met His His His His
                                                 1               5
```

```
cat cac acg gcc gag tcc gat asc ttc mag ctg tcc cag ggt ggg cag  104
His His Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln
            10                  15                  20

ggc ttc gcc att ccg atc ggg cag gcg atg gcg atc gcg ggc cag atc  152
Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile
        25                  30                  35

ggc tcg ggt ggg ggg tca ccc acc gtt cat atc ggg ccc acc gcc ttc  200
Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe
    40                  45                  50

ctc ggc ttg ggt gtt gtc gac aac aac ggc aac ggc gca ggg gtc caa  248
Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln
55                  60                  65

ggc gtg gtc ggg agc gct ccg gcg gcm agt ccc ggc atc tcc acc ggc  296
Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly
70                  75                  80                  85

gac gag atc acc ggg gtc gac ggc gct ccg atc aac tcg gcc acc gcg  344
Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala
                90                  95                  100

atg gcg gac gcg ctt asc ggg cat cat ccc ggt gac gtc atc tcg gtg  392
Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val
            105                 110                 115

acc tgg cma acc aag tcg ggc ggt acg cgt ama ggg aac gtg tca ttg  440
Thr Trp Xaa Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu
        120                 125                 130

gcc gag ggc ccc ccg gcc gaa ttc atg gtg gat ttc ggg gcg tta gca  488
Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro
    135                 140                 145

ccg gag atc aac tcc gcg agg atg tac gcc ggc ccg ggt tcg gcc tcg  536
Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser
150                 155                 160                 165

ctg gtg gcc gcy gct cag atg tgg gac agc gtg gcg agc gac ctg ttt  584
Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe
                170                 175                 180

tcg gtc gcg tcg gcg ttt cag tcg gtg gtc tgg ggt atg acg gtg ggg  632
Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly
            185                 190                 195

tcg tgg atg ggt tcg tcg gcg ggt ctg atg gtg gcg gcg gcc tcg ccg  680
Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro
        200                 205                 210

tat gtg gcg tgg atg agc gtc acc gcg ggg cag gcc gag ctg acc gcc  728
Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala
    215                 220                 225

gcc cag gtc cgg gtt gct gcg gcg gcc tac gag acg gcg cat ggg atg  776
Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu
230                 235                 240                 245

acg gtg acc cag acg gtg atc gcc gag aac cgt gct gaa ctg atg att  824
Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile
                250                 255                 260

ctg atc gcg acc aac ctc ttg ggg cha aac acc ccg gcg atc gcg gtc  872
Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val
            265                 270                 275
```

```
aac gag gcc gac tac ggc gag atg tgg gcc cag gac gcc gcc gcg atg   930
Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met
        280                 285                 290

ttt ggc tac gcc gcg gcg acg gcg acg gcg acg gcg acg ttg ctg cag   978
Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Gln
    295                 300                 305

ttc gag gag gcg ccg gag atg acc agc gcg gat ggg ctc ctc gag cag  1026
Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln
310                 315                 320                 325

gcc gcg gcg gtc gag gag gcc tcc gac acc gcc gcg gcg aac cag ttg  1074
Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu
                330                 335                 340

atg aac aat gtg ccc cag gcg ctg caa cag ctg gcc cag ccc acg cag  1122
Met Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln
            345                 350                 355

ggc acc acg cct tct tcc aag ctg ggt ggc ctg tgg aag acg gtc tcg  1170
Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser
        360                 365                 370

ccg cat cgg tcg ccg atc agc aac atg gtg tcg aug gcc aac aac cac  1218
Pro His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His
    375                 380                 385

atg tcg atg acc aac tcg ggt gtg ccg atg acc acc acc ttg agc tcg  1266
Met Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser
390                 395                 400                 405

atg ttg aag ggc ttt ggt cag gcg ggg gcc cgc cag gcc gtg caa acc  1314
Met Leu Lys Gly Phe Gly Gln Ala Pro Ala Ala Ala Arg Gln Ala Val Gln Thr
                410                 415                 420

gcg gcg caa aac ggg gtc cag gcg atg agc tcg ctg ggc agc tcg ctg  1362
Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu
            425                 430                 435

ggt tct tcg ggt ctg ggc ggt ggg gtg gcc ggc aac ttg ggt cgg gcg  1410
Gly Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala
        440                 445                 450

gcc tcg gtc ggt tcg ttg acg gtg ccg cag gcc tgg gcc gcg gcc aac  1458
Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn
    455                 460                 465

cag gca gcc acc acc ccg gcg ccg cgc ctg ccg ctg acc agc ctg acc  1506
Gln Ala Ala Thr Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr
470                 475                 480                 485

agc gcc gcg gaa aga ggg ccc ggg cag atg ctg ggc ggg ctg ccg gtg  1554
Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val
                490                 495                 500

ggg cag atg ggc gcc agg gcc ggt ggt ggg ctc agt ggt gtg ctg cgt  1602
Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg
            505                 510                 515

gtt ccg ccg cga ccc tat gtg atg ccg cat tat ccg gca gcc ggc gat  1650
Val Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp
        520                 525                 530

atc gcc ccg ccg gcc ttg ccg cag gac cgg ttc gcc gac ttc ccc gcg  1698
Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala
    535                 540                 545
```

```
ctg ccc gtc gac ccg tcc gcg atg gtc gcc caa gtg ggg cca cag gtg   1736
Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val
550                 555                 560                 565

gtc aac atc aac acc aca ctg ggc tac aac aac gcc gtg ggc gcc ggg   1784
Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly
                570                 575                 580

acc ggc atc gtc atc gat ccc aac ggt gtc gtg ctg acc aac aac cac   1832
Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His
            585                 590                 595

gtg atc gcg ggc gcc acc gac atc aat gcg ttc agc gtc ggc tcc ggc   1880
Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly
        600                 605                 610

caa acc tac ggc gtc gat gtg gtc ggg tat gac cgc acc cag gat gtc   1928
Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val
    615                 620                 625

gcg gtg ctg cag ctg cgc ggt gcc ggc ggc ctg ccg tcg gcg gcg atc   1976
Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile
630                 635                 640                 645

ggt ggc ggc gtc gcg gtt ggt gag ccc gcc gtc gcg atg ggc aac agc   2024
Gly Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser
                650                 655                 660

ggt ggg cag ggc ggg acg acc cgt ggg gtg cct ggc agg gtg gtc gcg   2072
Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala
            665                 670                 675

ctc ggc caa acc gtg cag gcg tcg gat tcg ctg acc ggt gcc gaa gag   2120
Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu
        680                 685                 690

acc ttg aac ggg ttg atc cag ttc gat gcc gcg atc cag ccc ggt gat   2168
Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp
    695                 700                 705

tcg ggc ggg ccc gtc gtc aac ggc cta gga cag gtg gtc ggt atg aac   2216
Ser Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn
710                 715                 720                 725

acg gcc gcg tcc tcgggatcc atcaccctgg cggccgctcg agcatgcatc       2268
Thr Ala Ala Ser

gattgtaacga agcctgaaa                                            2287

<210> SEQ ID NO:16
<211> 729
<212> PRT
<213> Artificial Sequence
<220> Description of Artificial Sequence:tri-fusion
      protein MTB72F (Rai2-TSH3-Ra35 or Mtb32-Mtb39
      fusion)

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
 1               5                  10                  15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
            20                  25                  30

Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile
        35                  40                  45

Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn
```

```
          50                  55                  60

Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu
 65                  70                  75                  80

Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile
                 85                  90                  95

Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly
            100                 105                 110

Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr
        115                 120                 125

Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp
    130                 135                 140

Phe Gly Ala Leu Pro Pro Gln Ile Asn Ser Ala Arg Met Tyr Ala Gly
145                 150                 155                 160

Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val
                165                 170                 175

Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp
            180                 185                 190

Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val
        195                 200                 205

Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln
    210                 215                 220

Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu
225                 230                 235                 240

Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg
                245                 250                 255

Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr
            260                 265                 270

Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Glu
        275                 280                 285

Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr
    290                 295                 300

Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly
305                 310                 315                 320

Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala
                325                 330                 335

Ala Ala Asn Glu Leu Met Asn Asn Val Pro Gln Ala Leu Gln Gln Leu
            340                 345                 350

Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu
        355                 360                 365

Trp Lys Thr Val Ser Pro His Arg Ser Pro Xaa Ser Asn Met Val Ser
    370                 375                 380

Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val Ser Met Thr
385                 390                 395                 400

Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Arg
                405                 410                 415

Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser
```

```
                420                             425                             430

Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly Val Ala Ala
            435                         440                         445

Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala
    450                         455                         460

Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg Ala Leu Pro
465                     470                         475                     480

Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu
                    485                         490                     495

Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu
            500                         505                         510

Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met Pro His Ser
        515                         520                         525

Pro Ala Ala Gly Asp Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe
    530                         535                         540

His Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln
545                     550                         555                     560

Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn
                    565                         570                     575

Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val
            580                         585                         590

Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe
        595                         600                         605

Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp
    610                         615                         620

Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu
625                     630                         635                     640

Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val Val
                    645                         650                     655

Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro
            660                         665                         670

Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu
        675                         680                         685

Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala
    690                         695                         700

Ile Gln Pro Gly Asp Ser Gly Gly Pro Val Val Asn Gly Leu Gly Gln
705                     710                         715                     720

Val Val Gly Met Asn Thr Ala Ala Ser
                    725


<210> SEQ ID NO:17
<211> 2190
<212> DNA
<213> Mtb72FMutSA

atgcatcacc atcaccatca cacggccggc tccgacaact tccagctgtc ccagggtggg 60
cagggattcg ccattccgat cgggcaggcg atggcgatcg cgggccagat ccgatcgggt 120
gggggtcacc ccaccgttca catcgggact aacggcttcc tcggcttggg tgttgtcgac 180
```

```
aacaacggca acggcgcacg agtccaacgc gtggtcggga gcgctccggc ggcaagtccc 240
ggcatctcca ccggcgacgt gatcaccgcg gtcgacggcg ctccgatcaa ctcggccacc 300
gcgatggcgg acgcgcttaa cgggcatcat cccggtgacg tcatctcggt gacctggcaa 360
accaagtcgg gcggcacgcg tacagggaac gtgacattgg ccgagggacc cccggccgaa 420
ttcatggtgg atttcggggc gttaccaccg gagatcaact ccgcgaggat gtacgccggc 480
ccgggttcgg cctcgctggt ggccgcggct cagatgtggg acagcgtggc gagtgacctg 540
ttttcggccg cgtcggcgtt tcagtcggtg gtctggggtc tgacggtggg gtcgtggata 600
ggttcgtcgg cgggtctgat ggtggcggcg gcctcgccgt atgtggcgtg gatgagcgtc 660
accgcggggc aggccgagct gaccgccgcc caggtccggg ttgctgcggc ggcctacgag 720
acggcgtatg ggctgacggt gcccccgccg gtgatcgccg agaaccgtgc tgaactgatg 780
attctgatag cgaccaacct cttggggcaa aacacccccg cgatcgcggt caacgaggcc 840
gaatacggcg agatgtgggc ccaagacgcc gccgcgatgt ttggctacgc cgcggcgacg 900
gcgacggcga cggcgacgtt gctgccgttc gaggaggcgc cggagatgac cagcgcgggt 960
gggctcctcg agcaggccgc cgcggtcgag gaggcctccg acaccgccgc ggcgaaccag 1020
ttgatgaaca atgtgcccca ggcgctgcaa cagctggccc agcccacgca gggcaccacg 1080
ccttcttcca agctgggtgg cctgtggaag acggtctcgc cgcatcggtc gccgatcagc 1140
aacatggtgt cgatggccaa caaccacatg tcgatgacca actcgggtgt gtcgatgacc 1200
aacaccttga gctcgatgtt gaagggcttt gctccggcgg cggccgccca ggccgtgcaa 1260
accgcggcgc aaaacggggt ccgggcgatg agctcgctgg gcagctcgct gggttcttcg 1320
ggtctgggcg gtggggtggc cgccaacttg ggtcgggcgg cctcggtcgg ttcgttgtcg 1380
gtgccgcagg cctgggccgc ggccaaccag gcagtcaccc cggcggcgcg ggcgctgccg 1440
ctgaccagcc tgaccagcgc cgcggaaaga gggcccgggc agatgctggg cgggctgccg 1500
gtggggcaga tgggcgccag ggccggtggt gggctcagtg gtgtgctgcg tgttccgccg 1560
cgaccctatg tgatgccgca ttctccggcc gccggcgata tcgccccgcc ggccttgtcg 1620
caggaccggt tcgccgactt ccccgcgctg cccctcgacc cgtccgcgat ggtcgcccaa 1680
gtggggccac aggtggtcaa catcaacacc aaactgggct acaacaacgc cgtgggcgcc 1740
gggaccggca tcgtcatcga tcccaacggt gtcgtgctga ccaacaacca cgtgatcgcg 1800
ggcgccaccg acatcaatgc gttcagcgtc ggctccggcc aaacctacgg cgtcgatgtg 1860
gtcgggtatg accgcaccca ggatgtcgcg gtgctgcagc tgcgcggtgc cggtggcctg 1920
ccgtcggcgg cgatcggtgg cggcgtcgcg gttggtgagc ccgtcgtcgc gatgggcaac 1980
agcggtgggc agggcggaac gccccgtgcg gtgcctggca gggtggtcgc gctcggccaa 2040
accgtgcagg cgtcggattc gctgaccggt gccgaagaga cattgaacgg gttgatccag 2100
ttcgatgccg cgatccagcc cggtgatgcg ggcgggcccg tcgtcaacgg cctaggacag 2160
gtggtcggta tgaacacggc cgcgtcctag                                  2190
```

<210> SEQ ID NO:18
<211> 729
<212> PRT
<213> Mtb72FMutSA

```
Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
                    5                  10                  15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
                20                  25                  30

Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile
            35                  40                  45

Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn
    50                  55                  60

Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu
65                  70                  75                  80

Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile
                85                  90                  95

Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly
            100                 105                 110

Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr
        115                 120                 125

Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp
    130                 135                 140
```

```
Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly
145                 150                 155                 160

Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Glu Met Trp Asp Ser Val
                165                 170                 175

Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp
            180                 185                 190

Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val
        195                 200                 205

Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln
    210                 215                 220

Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu
225                 230                 235                 240

Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg
                245                 250                 255

Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr
            260                 265                 270

Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln
        275                 280                 285

Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr
    290                 295                 300

Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly
305                 310                 315                 320

Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala
                325                 330                 335

Ala Ala Asn Gln Leu Met Asn Asn Val Pro Glu Ala Leu Gln Gln Leu
            340                 345                 350

Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu
        355                 360                 365

Trp Lys Thr Val Ser Pro Ala Arg Ser Pro Ile Ser Asn Met Val Ser
    370                 375                 380

Met Ala Asn Asn Lys Met Ser Met Thr Asn Ser Gly Val Ser Met Thr
385                 390                 395                 400

Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala
                405                 410                 415

Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser
            420                 425                 430

Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Val Ala Ala
        435                 440                 445

Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala
    450                 455                 460

Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg Ala Leu Pro
465                 470                 475                 480

Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu
                485                 490                 495

Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu
            500                 505                 510
```